Ste

TH

OIPE

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/980,054A

DATE: 12/04/2002 P.6
TIME: 10:27:05

Input Set : A:\19624051.app
Output Set: N:\CRF4\12032002\I980054A.raw

```
    3  <110> APPLICANT: LALANNE, JEAN L
    4         ROCHER, CORINNE
    6  <120> TITLE OF INVENTION: Novel genes of Candida albicans and the proteins coded
    7         by these genes
    9  <130> FILE REFERENCE: 146.1374
   11  <140> CURRENT APPLICATION NUMBER: 09/980,054A
C-->12  <141> CURRENT FILING DATE: 2002-04-14
   14  <150> PRIOR APPLICATION NUMBER: FR 9907250
   15  <151> PRIOR FILING DATE: 1999-06-09
   17  <150> PRIOR APPLICATION NUMBER: PCT/FR00/01567
   18  <151> PRIOR FILING DATE: 2000-06-08
   20  <160> NUMBER OF SEQ ID NOS: 32
   22  <170> SOFTWARE: PatentIn Ver. 2.1
   24  <210> SEQ ID NO: 1
   25  <211> LENGTH: 747
   26  <212> TYPE: DNA
   27  <213> ORGANISM: Candida albicans
   29  <220> FEATURE:
   30  <221> NAME/KEY: CDS
   31  <222> LOCATION: (1)..(747)
   33  <220> FEATURE:
   34  <221> NAME/KEY: modified_base
   35  <222> LOCATION: (136)..(138)
   37  <400> SEQUENCE: 1
   38  atg tca aat gac gat ata ata ctc cca tca gtt tca tcc tta tcg aaa   48
   39  Met Ser Asn Asp Asp Ile Ile Leu Pro Ser Val Ser Ser Leu Ser Lys
   40   1               5                  10                  15
   42  cta act ata aat gat gta tca aaa tca gga ttt gga tac aat ccg tcc   96
   43  Leu Thr Ile Asn Asp Val Ser Lys Ser Gly Phe Gly Tyr Asn Pro Ser
   44               20                  25                  30
   46  ata gga cca ata tca aat act att acc cta gaa tct tca ctg gta tta  144
   47  Ile Gly Pro Ile Ser Asn Thr Ile Thr Leu Glu Ser Ser Leu Val Leu
   48           35                  40                  45
   50  tta aat aaa cgt aca ata tca tta aca cca aca tca tct gac tcc att  192
   51  Leu Asn Lys Arg Thr Ile Ser Leu Thr Pro Thr Ser Ser Asp Ser Ile
   52       50                  55                  60
   54  tat gat aga aat att atc acg aaa aag cca cac gaa atc aac tta tct  240
   55  Tyr Asp Arg Asn Ile Ile Thr Lys Lys Pro His Glu Ile Asn Leu Ser
   56   65                  70                  75                  80
   58  tcg tta tca ttt ttg ttt tgt gag att att agt tgg gca cac tct aat  288
   59  Ser Leu Ser Phe Leu Phe Cys Glu Ile Ile Ser Trp Ala His Ser Asn
   60               85                  90                  95
   62  tcc aaa ggc att caa gat tta gaa aat cgt tta aac gga tta ggt tat  336
```

RAW SEQUENCE LISTING
PATENT APPLICATION: **US/09/980,054A**

DATE: 12/04/2002
TIME: 10:27:05

Input Set : **A:\19624051.app**
Output Set: **N:\CRF4\12032002\I980054A.raw**

```
63 Ser Lys Gly Ile Gln Asp Leu Glu Asn Arg Leu Asn Gly Leu Gly Tyr
64           100                 105             110
66 caa ata ggt caa cga tat ctc gaa ttg tgt aaa ata aga gaa ggt ttt     384
67 Gln Ile Gly Gln Arg Tyr Leu Glu Leu Cys Lys Ile Arg Glu Gly Phe
68           115                 120                 125
70 aaa aac agt aaa cga gag att aga ctt ttg gaa atg tta caa ttt att     432
71 Lys Asn Ser Lys Arg Glu Ile Arg Leu Leu Glu Met Leu Gln Phe Ile
72           130                 135                 140
74 cat ggt ccg ttc tgg aaa ttg att ttt ggt aaa act gct aat gaa tta     480
75 His Gly Pro Phe Trp Lys Leu Ile Phe Gly Lys Thr Ala Asn Glu Leu
76 145                 150                 155                 160
78 gaa aaa tcg caa gat ttg ccc aat gaa tat atg att gtg gag aat gtg     528
79 Glu Lys Ser Gln Asp Leu Pro Asn Glu Tyr Met Ile Val Glu Asn Val
80                 165                 170                 175
82 cca tta tta aat aga ttt att agt ata cct aag gag tat ggc gac tta     576
83 Pro Leu Leu Asn Arg Phe Ile Ser Ile Pro Lys Glu Tyr Gly Asp Leu
84           180                 185                 190
86 aat tgt tca gca ttt gtt gcg ggt ata att gag gga gca ctt gat aat     624
87 Asn Cys Ser Ala Phe Val Ala Gly Ile Ile Glu Gly Ala Leu Asp Asn
88           195                 200                 205
90 agt gga ttc aat gcc gat gtt aca gca cac acg gtc gct aca gat gca     672
91 Ser Gly Phe Asn Ala Asp Val Thr Ala His Thr Val Ala Thr Asp Ala
92           210                 215                 220
94 aat cca tta aga aca gta ttt ttg atc aag ttt gac gat tct gtt tta     720
95 Asn Pro Leu Arg Thr Val Phe Leu Ile Lys Phe Asp Asp Ser Val Leu
96 225                 230                 235                 240
98 att aga gag agt ttg aga ttt gga taa                                 747
99 Ile Arg Glu Ser Leu Arg Phe Gly
100                 245
103 <210> SEQ ID NO: 2
104 <211> LENGTH: 248
105 <212> TYPE: PRT
106 <213> ORGANISM: Candida albicans
108 <400> SEQUENCE: 2
109 Met Ser Asn Asp Asp Ile Ile Leu Pro Ser Val Ser Ser Leu Ser Lys
110   1               5                 10                  15
111 Leu Thr Ile Asn Asp Val Ser Lys Ser Gly Phe Gly Tyr Asn Pro Ser
112           20                  25                  30
113 Ile Gly Pro Ile Ser Asn Thr Ile Thr Leu Glu Ser Ser Leu Val Leu
114           35                  40                  45
115 Leu Asn Lys Arg Thr Ile Ser Leu Thr Pro Thr Ser Ser Asp Ser Ile
116     50                  55                  60
117 Tyr Asp Arg Asn Ile Ile Thr Lys Lys Pro His Glu Ile Asn Leu Ser
118   65                  70                  75                  80
119 Ser Leu Ser Phe Leu Phe Cys Glu Ile Ile Ser Trp Ala His Ser Asn
120                 85                  90                  95
121 Ser Lys Gly Ile Gln Asp Leu Glu Asn Arg Leu Asn Gly Leu Gly Tyr
122           100                 105             110
123 Gln Ile Gly Gln Arg Tyr Leu Glu Leu Cys Lys Ile Arg Glu Gly Phe
```

RAW SEQUENCE LISTING DATE: 12/04/2002
PATENT APPLICATION: **US/09/980,054A** TIME: 10:27:05

Input Set : **A:\19624051.app**
Output Set: **N:\CRF4\12032002\I980054A.raw**

```
124         115                120                125
125 Lys Asn Ser Lys Arg Glu Ile Arg Leu Leu Glu Met Leu Gln Phe Ile
126     130                135                140
127 His Gly Pro Phe Trp Lys Leu Ile Phe Gly Lys Thr Ala Asn Glu Leu
128 145                150                155                160
129 Glu Lys Ser Gln Asp Leu Pro Asn Glu Tyr Met Ile Val Glu Asn Val
130             165                170                175
131 Pro Leu Leu Asn Arg Phe Ile Ser Ile Pro Lys Glu Tyr Gly Asp Leu
132         180                185                190
133 Asn Cys Ser Ala Phe Val Ala Gly Ile Ile Glu Gly Ala Leu Asp Asn
134         195                200                205
135 Ser Gly Phe Asn Ala Asp Val Thr Ala His Thr Val Ala Thr Asp Ala
136     210                215                220
137 Asn Pro Leu Arg Thr Val Phe Leu Ile Lys Phe Asp Asp Ser Val Leu
138 225                230                235                240
139 Ile Arg Glu Ser Leu Arg Phe Gly
140                 245
144 <210> SEQ ID NO: 3
145 <211> LENGTH: 711
146 <212> TYPE: DNA
147 <213> ORGANISM: Candida albicans
149 <220> FEATURE:
150 <221> NAME/KEY: CDS
151 <222> LOCATION: (1)..(711)
153 <220> FEATURE:
154 <221> NAME/KEY: modified_base
155 <222> LOCATION: (577)..(579)
157 <400> SEQUENCE: 3
158 atg gat att gac gat att tta aaa gaa ttt gaa gag tct tca aaa gat    48
159 Met Asp Ile Asp Asp Ile Leu Lys Glu Phe Glu Glu Ser Ser Lys Asp
160   1           5                10                15
162 gaa aag att agc agt aaa aca tcg tct atc aac tta tat caa gac ttg    96
163 Glu Lys Ile Ser Ser Lys Thr Ser Ser Ile Asn Leu Tyr Gln Asp Leu
164         20                25                30
166 cta aga gct atg atc aac gaa cgt atg gct ccg gaa tta ttg cca tac   144
167 Leu Arg Ala Met Ile Asn Glu Arg Met Ala Pro Glu Leu Leu Pro Tyr
168         35                40                45
170 aaa caa gat tta atg tcc act gtt tta aca atg atg tct aac caa caa   192
171 Lys Gln Asp Leu Met Ser Thr Val Leu Thr Met Met Ser Asn Gln Gln
172     50                55                60
174 caa tat tta tta gaa tct cac gaa tat ggt gat atg aat ggc gac agt   240
175 Gln Tyr Leu Leu Glu Ser His Glu Tyr Gly Asp Met Asn Gly Asp Ser
176 65                70                75                80
178 ggt gta tta tcc gga gac ttt aaa tta caa cta atg att atc gaa act   288
179 Gly Val Leu Ser Gly Asp Phe Lys Leu Gln Leu Met Ile Ile Glu Thr
180             85                90                95
182 gat tta gag cgt ctc aac tat att gtt cga tta tac ata cga act cga   336
183 Asp Leu Glu Arg Leu Asn Tyr Ile Val Arg Leu Tyr Ile Arg Thr Arg
184         100                105                110
```

RAW SEQUENCE LISTING
PATENT APPLICATION: **US/09/980,054A**

DATE: 12/04/2002
TIME: 10:27:05

Input Set : **A:\19624051.app**
Output Set: **N:\CRF4\12032002\I980054A.raw**

```
186 ttg agt aag ttg aat aaa ttt act att ttt tac atc aat gaa agc agt   384
187 Leu Ser Lys Leu Asn Lys Phe Thr Ile Phe Tyr Ile Asn Glu Ser Ser
188         115                 120                 125
190 caa aat gat aat tta ttg tcc aaa gag gaa aga gat tat ata cac aaa   432
191 Gln Asn Asp Asn Leu Leu Ser Lys Glu Glu Arg Asp Tyr Ile His Lys
192         130                 135                 140
194 tat ttc cag att ttg act caa tta tat aac aac tgt ttc ctc aaa aaa   480
195 Tyr Phe Gln Ile Leu Thr Gln Leu Tyr Asn Asn Cys Phe Leu Lys Lys
196 145                 150                 155                 160
198 cta cca caa atg ttg acc tat ttg gat gac acc agt ggt gga caa tca   528
199 Leu Pro Gln Met Leu Thr Tyr Leu Asp Asp Thr Ser Gly Gly Gln Ser
200                 165                 170                 175
202 atg atc gtt gag cca gat tta gac cag cct gtg ttt atc aaa tgt acc   576
203 Met Ile Val Glu Pro Asp Leu Asp Gln Pro Val Phe Ile Lys Cys Thr
204                 180                 185                 190
206 ctg gaa gtc cca ata tta cta gat tac gac ggt gct aca gag ata gat   624
207 Leu Glu Val Pro Ile Leu Leu Asp Tyr Asp Gly Ala Thr Glu Ile Asp
208                 195                 200                 205
210 tta gaa tta ata aaa aag gga gtc tac gtg gtg aaa tac agc cta gtc   672
211 Leu Glu Leu Ile Lys Lys Gly Val Tyr Val Val Lys Tyr Ser Leu Val
212         210                 215                 220
214 aaa aga tat att gat att gga gat gtg gta ttg ata tga           711
215 Lys Arg Tyr Ile Asp Ile Gly Asp Val Val Leu Ile
216 225                 230                 235
219 <210> SEQ ID NO: 4
220 <211> LENGTH: 236
221 <212> TYPE: PRT
222 <213> ORGANISM: Candida albicans
224 <400> SEQUENCE: 4
225 Met Asp Ile Asp Asp Ile Leu Lys Glu Phe Glu Glu Ser Ser Lys Asp
226   1                 5                 10                  15
227 Glu Lys Ile Ser Ser Lys Thr Ser Ser Ile Asn Leu Tyr Gln Asp Leu
228                 20                  25                  30
229 Leu Arg Ala Met Ile Asn Glu Arg Met Ala Pro Glu Leu Leu Pro Tyr
230                 35                  40                  45
231 Lys Gln Asp Leu Met Ser Thr Val Leu Thr Met Met Ser Asn Gln Gln
232         50                  55                  60
233 Gln Tyr Leu Leu Glu Ser His Glu Tyr Gly Asp Met Asn Gly Asp Ser
234 65                  70                  75                  80
235 Gly Val Leu Ser Gly Asp Phe Lys Leu Gln Leu Met Ile Ile Glu Thr
236                 85                  90                  95
237 Asp Leu Glu Arg Leu Asn Tyr Ile Val Arg Leu Tyr Ile Arg Thr Arg
238                 100                 105                 110
239 Leu Ser Lys Leu Asn Lys Phe Thr Ile Phe Tyr Ile Asn Glu Ser Ser
240         115                 120                 125
241 Gln Asn Asp Asn Leu Leu Ser Lys Glu Glu Arg Asp Tyr Ile His Lys
242         130                 135                 140
243 Tyr Phe Gln Ile Leu Thr Gln Leu Tyr Asn Asn Cys Phe Leu Lys Lys
244 145                 150                 155                 160
```

```
245 Leu Pro Gln Met Leu Thr Tyr Leu Asp Asp Thr Ser Gly Gly Gln Ser
246               165                170                175
247 Met Ile Val Glu Pro Asp Leu Asp Gln Pro Val Phe Ile Lys Cys Thr
248               180                185                190
249 Leu Glu Val Pro Ile Leu Leu Asp Tyr Asp Gly Ala Thr Glu Ile Asp
250           195                200                205
251 Leu Glu Leu Ile Lys Lys Gly Val Tyr Val Val Lys Tyr Ser Leu Val
252       210                215                220
253 Lys Arg Tyr Ile Asp Ile Gly Asp Val Val Leu Ile
254 225           .   230                235
258 <210> SEQ ID NO: 5
259 <211> LENGTH: 1383
260 <212> TYPE: DNA
261 <213> ORGANISM: Candida albicans
263 <220> FEATURE:
264 <221> NAME/KEY: CDS
265 <222> LOCATION: (1)..(1383)
267 <400> SEQUENCE: 5
268 atg gat ttc ata gga gag att ata gag cat gag aca gag gca cct aaa    48
269 Met Asp Phe Ile Gly Glu Ile Ile Glu His Glu Thr Glu Ala Pro Lys
270  1               5                10                15
272 gaa cca acc cca aaa ccc aca att ggt gga ttc ccc gaa ctt aaa aaa    96
273 Glu Pro Thr Pro Lys Pro Thr Ile Gly Gly Phe Pro Glu Leu Lys Lys
274               20                25                30
276 tta aaa gaa aag aaa gtc tca aga tgg agg caa aag caa caa cag gaa   144
277 Leu Lys Glu Lys Lys Val Ser Arg Trp Arg Gln Lys Gln Gln Gln Glu
278           35                40                45
280 cag agc aca act tcc cca aaa act act gaa atc cgt tca gag gct tcc   192
281 Gln Ser Thr Thr Ser Pro Lys Thr Thr Glu Ile Arg Ser Glu Ala Ser
282       50                55                60
284 aaa att cac caa gaa aat atc gag aag atg gct caa atg tca gag gaa   240
285 Lys Ile His Gln Glu Asn Ile Glu Lys Met Ala Gln Met Ser Glu Glu
286  65                70                75                80
288 gag att ttg caa gag cgt gag gag tta cta aag ggt tta gat cct aaa   288
289 Glu Ile Leu Gln Glu Arg Glu Glu Leu Leu Lys Gly Leu Asp Pro Lys
290               85                90                95
292 tta att gaa agt ttg att ggt aga tcc aag aaa agg gaa gca aca gac   336
293 Leu Ile Glu Ser Leu Ile Gly Arg Ser Lys Lys Arg Glu Ala Thr Asp
294               100               105               110
296 cat gaa cac aat gga cat gct cat gaa cat gca gag gga tac cat gga   384
297 His Glu His Asn Gly His Ala His Glu His Ala Glu Gly Tyr His Gly
298           115               120               125
300 tgg att gga tca atg aaa act tct gaa gga tta aca gat tta tct caa   432
301 Trp Ile Gly Ser Met Lys Thr Ser Glu Gly Leu Thr Asp Leu Ser Gln
302       130               135               140
304 tta gat aag gaa gat gtg gac cgt gca ttg ggt ata agt tca tta tcc   480
305 Leu Asp Lys Glu Asp Val Asp Arg Ala Leu Gly Ile Ser Ser Leu Ser
306 145               150               155               160
308 tta tct gaa cct gag ggt ggc agt aat acg aaa aaa gtc gct ttc gac   528
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 12/04/2002
PATENT APPLICATION: US/09/980,054A   TIME: 10:27:06

Input Set : A:\19624051.app
Output Set: N:\CRF4\12032002\I980054A.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:21; N Pos. 18

VERIFICATION SUMMARY   DATE: 12/04/2002
PATENT APPLICATION: **US/09/980,054A**   TIME: 10:27:06

Input Set : **A:\19624051.app**
Output Set: **N:\CRF4\12032002\I980054A.raw**

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1308 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21 after pos.:0